



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant : Jan Zavada et al.

Serial No.: 09/967,237

Group Art Unit: 1642

Filed : September 27, 2001

Examiner: David J. Blanchard

For : MN Gene and Protein

SUBMISSION OF SUBSTITUTE SEQUENCE LISTING

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, Va 22313-1450

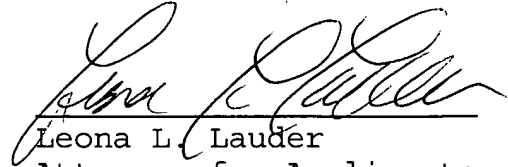
Sir:

Applicants submit the enclosed substitute Sequence Listing of the nucleotide and amino acid sequences contained in the above-identified application. Also enclosed is a computer readable copy of the substitute Sequence Listing. The nucleotide and amino acid sequences are presented in a form which conforms with the requirements of 37 CFR Sections 1.821 through 1.825.

In accordance with 37 CFR Section 1.821(f), the undersigned Attorney for the Applicants hereby states that the information recorded in computer readable form is identical to that in the printed substitute Sequence Listing. Further, in accordance with 37 CFR Section 1.821(g), the undersigned Attorney

for the Applicants states that the enclosed substitute Sequence Listing includes no new matter.

Respectfully submitted,

A handwritten signature in black ink, appearing to read 'Leona L. Lauder', is written over a horizontal line.

Leona L. Lauder
Attorney for Applicants
Registration No. 30,863

Dated: July 22, 2004



SEQUENCE LISTING

<110> Zavada, Jan
Pastorekova, Silvia
Pastorek, Jaromir

<120> MN Gene and Protein

<130> D-0021.5B-2

<140> 09/967,237

<141> 2001-09-27

<150> 09/178,115

<151> 1998-10-23

<160> 116

<170> PatentIn Ver. 2.1

<210> 1

<211> 1522

<212> DNA

<213> HUMAN

<220>

<221> CDS

<222> (13)..(1389)

<220>

<221> mat_peptide

<222> (124)..(1389)

<400> 1

acagtcagcc gc atg gct ccc ctg tgc ccc agc ccc tgg ctc cct ctg ttg 51
Met Ala Pro Leu Cys Pro Ser Pro Trp Leu Pro Leu Leu
-35 -30 -25

atc ccg gcc cct gct cca ggc ctc act gtg caa ctg ctg ctg tca ctg 99
Ile Pro Ala Pro Ala Pro Gly Leu Thr Val Gln Leu Leu Leu Ser Leu
-20 -15 -10

ctg ctt ctg atg cct gtc cat ccc cag agg ttg ccc cgg atg cag gag 147
Leu Leu Leu Met Pro Val His Pro Gln Arg Leu Pro Arg Met Gln Glu
-5 -1 1 5

gat tcc ccc ttg gga gga ggc tct tct ggg gaa gat gac cca ctg ggc 195
Asp Ser Pro Leu Gly Gly Gly Ser Ser Gly Glu Asp Asp Pro Leu Gly
10 15 20

gag gag gat ctg ccc agt gaa gag gat tca ccc aga gag gag gat cca 243
Glu Glu Asp Leu Pro Ser Glu Glu Asp Ser Pro Arg Glu Glu Asp Pro
25 30 35 40

ccc gga gag gag gat cta cct gga gag gag gat cta cct gga gag gag 291
Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu
45 50 55

gat cta cct gaa gtt aag cct aaa tca gaa gaa gag ggc tcc ctg aag	339
Asp Leu Pro Glu Val Lys Pro Lys Ser Glu Glu Glu Gly Ser Leu Lys	
60 65 70	
tta gag gat cta cct act gtt gag gct cct gga gat cct caa gaa ccc	387
Leu Glu Asp Leu Pro Thr Val Glu Ala Pro Gly Asp Pro Gln Glu Pro	
75 80 85	
cag aat aat gcc cac agg gac aaa gaa ggg gat gac cag agt cat tgg	435
Gln Asn Asn Ala His Arg Asp Lys Glu Gly Asp Asp Gln Ser His Trp	
90 95 100	
cgc tat gga ggc gac ccg ccc tgg ccc cgg gtg tcc cca gcc tgc gcg	483
Arg Tyr Gly Gly Asp Pro Pro Trp Pro Arg Val Ser Pro Ala Cys Ala	
105 110 115 120	
ggc cgc ttc cag tcc ccg gtg gat atc cgc ccc cag ctc gcc gcc ttc	531
Gly Arg Phe Gln Ser Pro Val Asp Ile Arg Pro Gln Leu Ala Ala Phe	
125 130 135	
tgc ccg gcc ctg cgc ccc ctg gaa ctc ctg ggc ttc cag ctc ccg ccg	579
Cys Pro Ala Leu Arg Pro Leu Glu Leu Leu Gly Phe Gln Leu Pro Pro	
140 145 150	
ctc cca gaa ctg cgc ctg cgc aac aat ggc cac agt gtg caa ctg acc	627
Leu Pro Glu Leu Arg Leu Arg Asn Asn Gly His Ser Val Gln Leu Thr	
155 160 165	
ctg cct cct ggg cta gag atg gct ctg ggt ccc ggg cgg gag tac cgg	675
Leu Pro Pro Gly Leu Glu Met Ala Leu Gly Pro Gly Arg Glu Tyr Arg	
170 175 180	
gct ctg cag ctg cat ctg cac tgg ggg gct gca ggt cgt ccg ggc tcg	723
Ala Leu Gln Leu His Leu His Trp Gly Ala Ala Gly Arg Pro Gly Ser	
185 190 195 200	
gag cac act gtg gaa ggc cac cgt ttc cct gcc gag atc cac gtg gtt	771
Glu His Thr Val Glu Gly His Arg Phe Pro Ala Glu Ile His Val Val	
205 210 215	
cac ctc agc acc gcc ttt gcc aga gtt gac gag gcc ttg ggg cgc ccg	819
His Leu Ser Thr Ala Phe Ala Arg Val Asp Glu Ala Leu Gly Arg Pro	
220 225 230	
gga ggc ctg gcc gtg ttg gcc gcc ttt ctg gag gag ggc ccg gaa gaa	867
Gly Gly Leu Ala Val Leu Ala Ala Phe Leu Glu Glu Gly Pro Glu Glu	
235 240 245	
aac agt gcc tat gag cag ttg ctg tct cgc ttg gaa gaa atc gct gag	915
Asn Ser Ala Tyr Glu Gln Leu Leu Ser Arg Leu Glu Glu Ile Ala Glu	
250 255 260	
gaa ggc tca gag act cag gtc cca gga ctg gac ata tct gca ctc ctg	963
Glu Gly Ser Glu Thr Gln Val Pro Gly Leu Asp Ile Ser Ala Leu Leu	
265 270 275 280	

ccc tct gac ttc agc cgc tac ttc caa tat gag ggg tct ctg act aca 1011
 Pro Ser Asp Phe Ser Arg Tyr Phe Gln Tyr Glu Gly Ser Leu Thr Thr
 285 290 295

ccg ccc tgt gcc cag ggt gtc atc tgg act gtg ttt aac cag aca gtg 1059
 Pro Pro Cys Ala Gln Gly Val Ile Trp Thr Val Phe Asn Gln Thr Val
 300 305 310

atg ctg agt gct aag cag ctc cac acc ctc tct gac acc ctg tgg gga 1107
 Met Leu Ser Ala Lys Gln Leu His Thr Leu Ser Asp Thr Leu Trp Gly
 315 320 325

cct ggt gac tct cgg cta cag ctg aac ttc cga gcg acg cag cct ttg 1155
 Pro Gly Asp Ser Arg Leu Gln Leu Asn Phe Arg Ala Thr Gln Pro Leu
 330 335 340

aat ggg cga gtg att gag gcc tcc ttc cct gct gga gtg gac agc agt 1203
 Asn Gly Arg Val Ile Glu Ala Ser Phe Pro Ala Gly Val Asp Ser Ser
 345 350 355 360

cct cgg gct gct gag cca gtc cag ctg aat tcc tgc ctg gct gct ggt 1251
 Pro Arg Ala Ala Glu Pro Val Gln Leu Asn Ser Cys Leu Ala Ala Gly
 365 370 375

gac atc cta gcc ctg gtt ttt ggc ctc ctt ttt gct gtc acc agc gtc 1299
 Asp Ile Leu Ala Leu Val Phe Gly Leu Leu Phe Ala Val Thr Ser Val
 380 385 390

gcg ttc ctt gtg cag atg aga agg cag cac aga agg gga acc aaa ggg 1347
 Ala Phe Leu Val Gln Met Arg Arg Gln His Arg Arg Gly Thr Lys Gly
 395 400 405

ggt gtg agc tac cgc cca gca gag gta gcc gag act gga gcc 1389
 Gly Val Ser Tyr Arg Pro Ala Glu Val Ala Glu Thr Gly Ala
 410 415 420

tagaggctgg atcttgaga atgtgagaag ccagccagag gcattctgagg gggagccggt 1449

aactgtcctg tctgtctcat tatgccactt ccttttaact gccaaagaat tttttaaaat 1509

aaatatttat aat 1522

<210> 2
 <211> 459
 <212> PRT
 <213> HUMAN

<400> 2
 Met Ala Pro Leu Cys Pro Ser Pro Trp Leu Pro Leu Leu Ile Pro Ala
 -35 -30 -25

Pro Ala Pro Gly Leu Thr Val Gln Leu Leu Leu Ser Leu Leu Leu Leu
 -20 -15 -10

Met Pro Val His Pro Gln Arg Leu Pro Arg Met Gln Glu Asp Ser Pro
 -5 -1 1 5 10

Leu Gly Gly Gly Ser Ser Gly Glu Asp Asp Pro Leu Gly Glu Glu Asp
 15 20 25
 Leu Pro Ser Ser Glu Glu Asp Ser Pro Arg Glu Glu Asp Pro Pro Gly Glu
 30 35 40
 Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro
 45 50 55
 Glu Val Lys Pro Lys Ser Glu Glu Glu Gly Ser Leu Lys Leu Glu Asp
 60 65 70 75
 Leu Pro Thr Val Glu Ala Pro Gly Asp Pro Gln Glu Pro Gln Asn Asn
 80 85 90
 Ala His Arg Asp Lys Glu Gly Asp Asp Gln Ser His Trp Arg Tyr Gly
 95 100 105
 Gly Asp Pro Pro Trp Pro Arg Val Ser Pro Ala Cys Ala Gly Arg Phe
 110 115 120
 Gln Ser Pro Val Asp Ile Arg Pro Gln Leu Ala Ala Phe Cys Pro Ala
 125 130 135
 Leu Arg Pro Leu Glu Leu Leu Gly Phe Gln Leu Pro Pro Leu Pro Glu
 140 145 150 155
 Leu Arg Leu Arg Asn Asn Gly His Ser Val Gln Leu Thr Leu Pro Pro
 160 165 170
 Gly Leu Glu Met Ala Leu Gly Pro Gly Arg Glu Tyr Arg Ala Leu Gln
 175 180 185
 Leu His Leu His Trp Gly Ala Ala Gly Arg Pro Gly Ser Glu His Thr
 190 195 200
 Val Glu Gly His Arg Phe Pro Ala Glu Ile His Val Val His Leu Ser
 205 210 215
 Thr Ala Phe Ala Arg Val Asp Glu Ala Leu Gly Arg Pro Gly Gly Leu
 220 225 230 235
 Ala Val Leu Ala Ala Phe Leu Glu Glu Gly Pro Glu Glu Asn Ser Ala
 240 245 250
 Tyr Glu Gln Leu Leu Ser Arg Leu Glu Glu Ile Ala Glu Glu Gly Ser
 255 260 265
 Glu Thr Gln Val Pro Gly Leu Asp Ile Ser Ala Leu Leu Pro Ser Asp
 270 275 280
 Phe Ser Arg Tyr Phe Gln Tyr Glu Gly Ser Leu Thr Thr Pro Pro Cys
 285 290 295
 Ala Gln Gly Val Ile Trp Thr Val Phe Asn Gln Thr Val Met Leu Ser
 300 305 310 315

Ala Lys Gln Leu His Thr Leu Ser Asp Thr Leu Trp Gly Pro Gly Asp
320 325 330

Ser Arg Leu Gln Leu Asn Phe Arg Ala Thr Gln Pro Leu Asn Gly Arg
335 340 345

Val Ile Glu Ala Ser Phe Pro Ala Gly Val Asp Ser Ser Pro Arg Ala
350 355 360

Ala Glu Pro Val Gln Leu Asn Ser Cys Leu Ala Ala Gly Asp Ile Leu
365 370 375

Ala Leu Val Phe Gly Leu Leu Phe Ala Val Thr Ser Val Ala Phe Leu
380 385 390 395

Val Gln Met Arg Arg Gln His Arg Arg Gly Thr Lys Gly Gly Val Ser
400 405 410

Tyr Arg Pro Ala Glu Val Ala Glu Thr Gly Ala
415 420

<210> 3
<211> 29
<212> DNA
<213> HUMAN

<400> 3
cgcccagtg gtcattctcc ccagaagag

29

<210> 4
<211> 19
<212> DNA
<213> HUMAN

<400> 4
ggaatcctcc tgcattccgg

19

<210> 5
<211> 10898
<212> DNA
<213> HUMAN

<220>
<221> gene
<222> (1)..(10898)
<223> full-length MN genomic sequence

<220>
<221> unsure of base at position 1974
<222> (1974)
<223> unsure of base at position 1974, which is in the 5' region flanking the transcription initiation site (3507) as determined by RNase protection assay.

<400> 5

```
ggatcctggt gactcgtgac cttaccccca accctgtgct ctctgaaaca tgagctgtgt 60
ccactcaggg ttaaattgat taagggcggt gcaagatgtg ctttggtaaa cagatgcttg 120
aaggcagcat gctcgttaag agtcatcacc aatccctaata ctcaagtaat cagggacaca 180
aacactgcgg aaggccgcag ggctcctctgc ctaggaaaac cagagacctt tgttcacttg 240
tttatctgac cttccctcca ctattgtcca tgaccctgcc aaatccccct ctgtgagaaa 300
cacccaagaa ttatcaataa aaaaataaat ttaaaaaaaa aatacaaaaa aaaaaaaaaa 360
aaaaaaaaaa gacttacgaa tagttattga taaatgaata gctattggta aagccaagta 420
aatgatcata ttcaaaacca gacggccatc atcacagctc aagtctacct gatttgatct 480
ctttatcatt gtcattcttt ggattcacta gattagtcac catcctcaaa attctcccc 540
aagttctaata tacgttccaa acatttaggg gttacatgaa gcttgaacct actaccttct 600
ttgcttttga gccatgagtt gtaggatgaa tgagtttaca ccttacctgc tggggattaa 660
tttaaacttt acctctaagt cagttgggtg gcctttggct tatttttcta gctaattttg 720
tagttaatgg atgcactgtg aatcttgcta tgatagtttt cctccacact ttgccactag 780
gggtaggtag gtactcagtt ttcagtaatt gcttacctaa gaccctaagc cctattttctc 840
ttgtactggc ctttatctgt aatatgggca tatttaatac aatataattt ttggagtttt 900
tttgtttgtt tgtttgtttg tttttttgag acggagtctt gcactctgtc tgcccaggct 960
ggagtagcag tgggtccatc tcggctcact gcaagctcca cctcccgagt tcacgccatt 1020
ttcctgcctc agcctcccga gtagctggga ctacaggcgc ccgccaccat gcccggttaa 1080
ttttttgtat ttttggtaga gacgggggtt caccgtgtta gccagaatgg tctcgatctc 1140
ctgacttcgt gatccaccgc cctcggcctc ccaaagttct gggattacag gtgtgagcca 1200
ccgcacctgg ccaatttttt gagtctttta aagtaaaaaat atgtcttgta agctggtaac 1260
tatggtacat ttccttttat taatgtgggt ctgacgggtc tataggttct tttgagtttg 1320
gcatgcatat gctacttttt gcagtccttt cattacattt ttctctcttc atttgaagag 1380
catgttatat ctttttagctt cacttggttt aaaaggttct ctcatagacc taacacagtg 1440
tcattgttgg taccacttgg atcataagtg gaaaaacagt caagaaattg cacagtaata 1500
cttgttttga agagggatga ttcaggtgaa tctgacacta agaaactccc ctacctgagg 1560
tctgagattc ctctgacatt gctgtatata ggcttttctt ttgacagcct gtgactgcgg 1620
actatttttc ttaagcaaga tatgctaaag ttttgtgagc ctttttccag agagaggtct 1680
catatctgca tcaagtgaga acatataatg tctgcatggt tccatatttc aggaatgttt 1740
gcttgtgttt tatgttttta tatagacagg gaaacttggt cctcagtgac ccaaaagagg 1800
tggaatttgt tattggatat catcattggc ccacgctttc tgaccttggg aacaattaag 1860
gggtcataat ctcaattctg tcagaattgg tacaagaaat agctgctatg tttcttgaca 1920
ttccacttgg taggaaataa gaatgtgaaa ctcttcagtt ggtgtgtgtc cctngttttt 1980
ttgcaatttc cttcttactg tgttaaaaaa aagtatgatc ttgctctgag aggtgaggca 2040
ttcttaatac tgatctttta agatcaataa tataatcctt tcaaggatta tgtctttatt 2100
ataataaaga taatttgtct ttaacagaat caataatata atcccttaaa ggattatatc 2160
tttgctgggc gcagtggctc acacctgtaa tcccagcact ttgggtggcc aaggtggaag 2220
gatcaaatth gcctacttct atattatctt cttaaagcaga attcatctct cttccctcaa 2280
tatgatgata ttgacagggt ttgccctcac tctactagatt gtgagctcct gctcagggca 2340
ggtagcgttt tttgtttttg tttttgtttt tcttttttga gacagggctt tgctctgtca 2400
cccaggccag agtgcaatgg tacagtctca gctcactgca gcctcaagcg cctcggtcca 2460
aaccatcatc ccatttcagc ctcttgagta gctgggacta caggcacatg ccattacacc 2520
tggtcaattt ttttgtattt ctagtagaga cagggttttg ccatgttgcc cgggctggtc 2580
tcgaactcct ggactcaagc aatccaccca cctcagcctc ccaaaatgag ggaccgtgtc 2640
ttattcattt ccattgctct agtccatagc ccagtgtctg acctatggta gtactaaata 2700
aatatttgtt gaattgcaata gtaaatagca tttcaggagg caagaactag attaacaaga 2760
gtggtaaaag gtttgagaaa aaaaataata gtttaatttg gctagagtat gagggagagt 2820
agtaggagac aagatggaaa ggtctcttgg gcaaggtttt gaaggaggtt ggaagtcaga 2880
agtacacaat agcatatctg tggcaggcag tggggagcca atgaaggctt ttgagcagga 2940
gagtaatgtg ttgaaaaata aatataggtt aaacctatca gagccctctt gacacataca 3000
cttgcctttc attcaagctc aagtttgtct ccacataacc cattacttaa ctcaccctcg 3060
ggctccctca gcagcctgcc ctacctcttt acctgcttcc tgggtggagt agggatgtat 3120
acatgagctg ctttccctct cagccagagg acatgggggg cccagctccc cctgcctttc 3180
cccttctgtg cctggagctg ggaagcaggc cagggttagc tgaggctggc tggcaagcag 3240
ctgggtgggt ccaggagag cctgcatagt gccaggtggg gccttgggtt ccaagctagt 3300
```


ccatggcccc	gataaccttc	tgctgtgca	cacacctgcc	cctcactcca	ccccatcct	3360
agcttttgta	tgggggagag	ggcacagggc	cagacaaacc	tgtgagactt	tggctccatc	3420
tctgcaaaag	ggcgctctgt	gagtcagcct	gctccccctc	aggcttgctc	ctccccacc	3480
cagctctcgt	ttccaatgca	cgtacagccc	gtacacaccg	tgtgctggga	caccccacag	3540
tcagccgcat	ggctccccctg	tgccccagcc	cctggctccc	tctgttgatc	ccggccccctg	3600
ctccaggcct	cactgtgcaa	ctgctgctgt	cactgctgct	tctgggtgct	gtccatcccc	3660
agaggttgcc	ccgatgagag	gaggattccc	ccttgggagg	aggctcttct	ggggaagatg	3720
accactggg	cgaggaggat	ctgcccagtg	aagaggattc	accagagag	gaggatccac	3780
ccggagagga	ggatctacct	ggagaggagg	atctacctgg	agaggaggat	ctacctgaag	3840
ttaagcctaa	atcagaagaa	gagggctccc	tgaagttaga	ggatctacct	actgttgagg	3900
ctcttgagag	tcctcaagaa	ccccagaata	atgccacag	ggacaaagaa	ggtaagtggg	3960
catcaatctc	caaatccagg	ttccaggagg	ttcatgactc	ccctcccata	ccccagccta	4020
ggctctgttc	actcagggaa	ggaggggaga	ctgtactccc	cacagaagcc	cttccaggag	4080
ttccatacca	atatccccat	ccccactctc	ggaggtagaa	agggacagat	gtggagagaa	4140
aataaaaagg	gtgcaaaagg	agagaggtga	gctggatgag	atgggagaga	agggggaggc	4200
tggagaagag	aaagggatga	gaactgcaga	tgagagaaaa	aatgtgcaga	cagaggaaaa	4260
aaataggtgg	agaaggagag	tcagagagtt	tgaggggaag	agaaaaggaa	agcttggggag	4320
gtgaagtggg	taccagagac	aagcaagaag	agctggtaga	agtcactctca	tcttaggcta	4380
caatgaggaa	ttgagaccta	ggaagaaggg	acacagcagg	tagagaaacg	tggcttcttg	4440
actccaagc	caggaatttg	gggaaagggg	ttggagacca	tacaaggcag	agggatgagt	4500
ggggagaaga	aagaaggag	aaaggaaaga	tggtgtactc	actcatttgg	gactcaggac	4560
tgaagtgcc	actcactttt	tttttttttt	tttttgagac	aaactttcac	ttttgttgcc	4620
caggctggag	tgcaatggcg	cgatctcggc	tcactgcaac	ctccacctcc	cgggttcaag	4680
tgattctcct	gcctcagcct	ctagccaagt	agctgcgatt	acaggcatgc	gccaccacgc	4740
ccggctaatt	tttgtatttt	tagtagagac	ggggtttcgc	catgttggtc	aggctggctc	4800
cgaactcctg	atctcaggtg	atccaaccac	cctggcctcc	caaagtgctg	ggattatagg	4860
cgtgagccac	agcgctggc	ctgaagcagc	cactcacttt	tacagaccct	aagacaatga	4920
ttgcaagctg	gtaggattgc	tgtttgcccc	accagctgc	ggtgttgagt	ttgggtgagg	4980
tctcctgtgc	tttgacctg	gcccgttaa	ggcatttggt	accgtaatg	ctcctgtaag	5040
gcatctgcgt	ttgtgacatc	gttttggtcg	ccaggaaggg	attggggctc	taagcttgag	5100
cggttcatcc	ttttcattta	tacaggggat	gaccagagtc	attggcgcta	tggaggtgag	5160
acaccacccc	cctgcacaga	cccaatctgg	gaaccagct	ctgtggatct	cccctacagc	5220
cgtccctgaa	gactgggtccc	gggcgtccca	cccgccgccc	accgtccac	cccctcacct	5280
tttctaccgg	ggttccctaa	gttctgacc	taggcgtcag	acttcctcac	tatactctcc	5340
caccccaggc	gacccgcctt	ggccccgggt	gtccccagcc	tgcgcgggccc	gcttccagtc	5400
ccgggtggat	atccgcccc	agctcgccgc	cttctgcccc	gccctgcgcc	ccctggaact	5460
cctgggcttc	cagctcccg	cgctcccaga	actgcgcctg	cgcaacaatg	gccacagtgg	5520
tgaggggggtc	tccccgccga	gacttgggga	tggggcgggg	cgcaggggaag	ggaaccgtcg	5580
cgcagtgcct	gccccggggg	tgggctggcc	ctaccgggcg	gggcgggctc	acttgcctct	5640
ccctacgcag	tgcaactgac	cctgcctcct	gggctagaga	tggctctggg	tcccgggcgg	5700
gagtaccggg	ctctgcagct	gcattctgcac	tggggggctg	caggctcgcc	gggctcggag	5760
cacactgtgg	aaggccaccg	tttccctgcc	gaggtgagcg	cggactggcc	gagaaggggc	5820
aaaggagcgg	ggcggacggg	ggccagagac	gtggccctct	cctaccctcg	tgtccttttc	5880
agatccacgt	ggttcacctc	agcaccgcct	ttgccagagt	tgacgaggcc	ttggggcgcc	5940
cgggaggcct	ggcgtggtg	gcccgttttc	tggaggtacc	agatcctgga	caccccctac	6000
tccccgcttt	cccattccat	gctcctcccg	gactctatcg	tggagccaga	gaccccatcc	6060
cagcaagctc	actcaggccc	ctggctgaca	aactcattca	cgcactgttt	gttcatttaa	6120
cacccactgt	gaaccaggca	ccagccccca	acaaggattc	tgaagctgta	ggtccttgcc	6180
tctaaggagc	ccacagccag	tgggggaggc	tgacatgaca	gacacatagg	aaggacatag	6240
taaagatggg	ggtcacagag	gaggtgacac	ttaaagcctt	cactggtaga	aaagaaaagg	6300
agggtttcat	tgagagagaa	acagaatgtg	caaagactca	gaatatggcc	tatttaggga	6360
atggctacat	acaccatgat	tagaggaggc	ccagtaaagg	gaagggatgg	tgagatgcct	6420
gctaggttca	ctcactcact	tttattttatt	tattttatttt	tttgacagtc	tctctgtcgc	6480
ccaggctgga	gtgcagtggg	gtgatcttgg	gtcactgcaa	cttccgcctc	ccgggttcaa	6540
gggattctcc	tgctcagct	tctgagtag	ctggggttac	agggtgtgtg	caccatgccc	6600
agctaatttt	tttttgattt	tttagtagac	agggtttcac	catgttggtc	aggctggctc	6660
caaactcctg	gcctcaagtg	atccgcctga	ctcagcctac	caaagtgtg	attacaagtg	6720

tgagccaccg	tgcccagcca	cactcactga	ttctttaatg	ccagccacac	agcacaaagt	6780
tcagagaaat	gcctccatca	tagcatgtca	atatgttcat	actcttaggt	tcattgatgtt	6840
cttaacatta	ggttcataag	caaaataaga	aaaaagaata	ataaataaaa	gaagtggcat	6900
gtcaggacct	cacctgaaaa	gccaaacaca	gaatcatgaa	ggtgaatgca	gaggtgacac	6960
caacacaaag	gtgtatatat	ggtttcctgt	ggggagtatg	tacggaggca	gcagtgagtg	7020
agactgcaaa	cgtcagaagg	gcacgggtca	ctgagagcct	agtatcctag	taaagtgggc	7080
tctctccctc	tctctccagc	ttgtcattga	aaaccagtc	accaagcttg	ttgggttcgca	7140
cagcaagagt	acatagagtt	tgaaataata	cataggattt	taagagggag	acactgtctc	7200
taaaaaaaaa	aacaacagca	acaacaaaaa	gcaacaacca	ttacaatttt	atgttccctc	7260
agcattctca	gagctgagga	atgggagagg	actatgggaa	cccccttcat	gttccggcct	7320
tcagccatgg	ccctggatac	atgcactcat	ctgtcttaca	atgtcattcc	cccaggaggg	7380
cccgaagaa	aacagtgcct	atgagcagtt	gctgtctcgc	ttggaagaaa	tcgctgagga	7440
aggtcagttt	ggttggtctgg	ccactaatct	ctgtggccta	gttcataaag	aatcaccctt	7500
tggagcttca	ggtctgaggc	tggagatggg	ctccctccag	tgcaggaggg	attgaagcat	7560
gagccagcgc	tcattctgat	aataaccatg	aagctgacag	acacagttac	ccgcaaacgg	7620
ctgcctacag	attgaaaacc	aagcaaaaac	cgccgggcac	ggtggctcac	gcctgtaatc	7680
ccagcacttt	gggaggccaa	ggcagggtga	tcacgaggtc	aagagatcaa	gaccatcctg	7740
gccaacatgg	tgaaacccca	tctctactaa	aaatacgaaa	aaatagccag	gcgtgggtggc	7800
gggtgcctgt	aatcccagct	actcgggagg	ctgaggcagg	agaatggcat	gaaccgggga	7860
ggcagaagtt	gcagtgagcc	gagatcgtgc	cactgcactc	cagcctgggc	aacagagcga	7920
gactcttgct	tcaaaaaaaaa	aaaaaaaaaaa	gaaaaccaag	caaaaaccaa	aatgagacaa	7980
aaaaaacaag	acaaaaaaat	ggtgttttga	aattgtcaag	gtcaagtctg	gagagctaaa	8040
ctttttctga	gaactgttta	tctttaataa	gcatacaata	ttttaacttt	gtaaataactt	8100
ttgttggaag	tcgttctctt	cttagtcact	cttgggtcat	tttaaactct	acttactcta	8160
ctagaccttt	taggtttctg	ctagactagg	tagaactctg	cctttgcatt	tcttgtgtct	8220
gttttgtata	gttatcaata	ttcatattta	tttacaagtt	attcagatca	ttttttcttt	8280
tctttttttt	tttttttttt	ttttttacat	cttttagtaga	gacagggttt	caccatattg	8340
gccaggtctc	tctcaaactc	ctgaccttgt	gatccaccag	cctcggcctc	ccaaagtgtc	8400
gggattcatt	ttttcttttt	aatttgctct	gggttaaac	ttgtggccca	gcactttatg	8460
atggtacaca	gagttaagag	tgtagactca	gacggctctt	cttctttcct	tctcttccct	8520
cctcccttcc	ctcccacett	cccttctctc	cttcttttct	ttcttctctc	cttgtctcct	8580
caggcctctt	ccagttgctc	caaagccctg	tacttttttt	tgagttaacg	tcttatggga	8640
agggcctgca	cttagtgaa	aagtgggtctc	agagttgagt	taccttggct	tctggggagt	8700
gaaactgtat	ccctataccc	tgaagcttta	agggggtgca	atgtagatga	gaccccaaca	8760
tagatcctct	tcacaggctc	agagactcag	gtcccaggac	tggacatatc	tgcactcctg	8820
ccctctgact	tcagccgcta	cttccaatat	gaggggtctc	tgactacacc	gccctgtgct	8880
cagggtgtca	tctggactgt	gtttaaccag	acagtgatgc	tgagtgtctaa	gcagggtgggc	8940
ctgggggtgtg	tgtggacaca	gtgggtgcgg	gggaaagagg	atgtaagatg	agatgagaaa	9000
caggagaaga	aagaaatcaa	ggctgggctc	tgtggcttac	gcctataatc	ccaccacgtt	9060
gggaggctga	ggtgggagaa	tggtttgagc	ccaggagttc	aagacaaggc	ggggcaacat	9120
agtgtgaccc	catctctacc	aaaaaaaaacc	caacaaaacc	aaaaatagcc	gggcatggtg	9180
gtatgcggcc	tagtcccagc	tactcaagga	ggctgaggtg	ggaagatcgc	ttgattccag	9240
gagtttgaga	ctgcagtgag	ctatgatccc	accactgcct	accatcttta	ggatacattt	9300
atattattat	aaaagaaatc	aagaggctgg	atggggaata	caggagctgg	aggggtggagc	9360
cctgagggtgc	tggttgtgag	ctggcctggg	acccttgttt	cctgtcatgc	catgaacca	9420
cccacactgt	ccactgacct	ccctagctcc	acaccctctc	tgacaccctg	tggggacctg	9480
gtgactctcg	gctacagctg	aacttccgag	cgacgcagcc	tttgaatggg	cgagtgattg	9540
aggcctcctt	ccctgctgga	gtggacagca	gtcctcgggc	tgtgtgagca	ggtacagctt	9600
tgtctggttt	ccccccagcc	agtagtccct	tatcctccca	tgtgtgtgcc	agtgtctgtc	9660
attggtggtc	acagcccgcc	tctcacatct	cctttttctc	tccagtccag	ctgaattcct	9720
gcctggctgc	tggtagtctc	gcccctcctc	ttggtcctga	tgccaggaga	ctcctcagca	9780
ccattcagcc	ccagggtctgc	tcaggaccgc	ctctgtctcc	tctccttttc	tgcagaacag	9840
accccaaccc	caatattaga	gaggcagatc	atggtgggga	ttccccattt	gtccccagag	9900
gctaattgat	tagaatgaag	cttgagaaat	ctcccagcat	ccctctcgca	aaagaatccc	9960
ccccctttt	tttaaagata	gggtctcact	ctgtttgccc	caggctgggg	tgttgtggca	10020
cgatcatagc	tactgcagc	ctcgaactcc	taggtcagg	caatcctttc	accttagctt	10080
ctcaaagcac	tgggactgta	ggcatgagcc	actgtgcctg	gccccaaacg	gccctttttac	10140

```

ttggcttttta ggaagcaaaa acggtgctta tcttaccctt tctcgtgtat ccaccctcat 10200
cccttggctg gcctcttctg gagactgagg cactatgggg ctgcctgaga actcggggca 10260
ggggtggtgg agtgactga ggcaggtgtt gaggaactct gcagaccctt cttccttccc 10320
aaagcagccc tctctgctct ccatcgcagg tgacatccta gccctggttt ttggcctcct 10380
ttttgctgtc accagcgtcg cgttccttgt gcagatgaga aggcagcaca ggtattacac 10440
tgaccctttc ttcaggcaca agcttcccc acccttgttg agtcacttca tgcaaagcgc 10500
atgcaaata gctgctcctg ggccagtttt ctgattagcc tttcctgttg tgtacacaca 10560
gaaggggaac caaagggggg gtgagctacc gcccagcaga ggtagccgag actggagcct 10620
agaggctgga tcttgagaa tgtgagaagc cagccagagg catctgaggg ggagccggtg 10680
actgtcctgt cctgtctcatt atgccacttc cttttaactg ccaagaaatt ttttaaaata 10740
aatatttata ataaaatatg tgtagtcac ctttggtccc caaatcagaa ggaggtattt 10800
gaatttccta ttactgttat tagcaccaat ttagtggtaa tgcatttatt ctattacagt 10860
tcggcctcct tccacacatc actccaatgt gttgctcc 10898

```

```

<210> 6
<211> 37
<212> PRT
<213> HUMAN

```

```

<400> 6
Met Ala Pro Leu Cys Pro Ser Pro Trp Leu Pro Leu Leu Ile Pro Ala
  1             5             10             15

```

```

Pro Ala Pro Gly Leu Thr Val Gln Leu Leu Leu Ser Leu Leu Leu Leu
      20             25             30

```

```

Met Pro Val His Pro
      35

```

```

<210> 7
<211> 25
<212> DNA
<213> HUMAN

```

```

<400> 7
tgggggttctt gaggatctcc aggag 25

```

```

<210> 8
<211> 26
<212> DNA
<213> HUMAN

```

```

<400> 8
ctctaacttc agggagccct cttctt 26

```

```

<210> 9
<211> 48
<212> DNA
<213> HUMAN

```

```

<220>
<221> primer_bind
<222> (1)..(48)

```

<223> anchor primer that anneals to the homopolymeric tail.

<220>

<221> inosine

<222> (36)..(37) (41)..(42) (46)..(47)

<223> each of the modified_bases at positions (36), (37), (41), (42), (46) and (47) are inosine

<400> 9

cuacuacuac uaggccacgc gtcgactagt acggaaggg aaggaag

48

<210> 10

<211> 6

<212> PRT

<213> HUMAN

<400> 10

Glu Glu Asp Leu Pro Ser

1

5

<210> 11

<211> 6

<212> PRT

<213> HUMAN

<400> 11

Gly Glu Asp Asp Pro Leu

1

5

<210> 12

<211> 21

<212> PRT

<213> HUMAN

<400> 12

Asn Asn Ala His Arg Asp Lys Glu Gly Asp Asp Gln Ser His Trp Arg

1

5

10

15

Tyr Gly Gly Asp Pro

20

<210> 13

<211> 16

<212> PRT

<213> HUMAN

<400> 13

His Pro Gln Arg Leu Pro Arg Met Gln Glu Asp Ser Pro Leu Gly Gly

1

5

10

15

<210> 14

<211> 24

<212> PRT
<213> HUMAN

<400> 14
Glu Glu Asp Ser Pro Arg Glu Glu Asp Pro Pro Gly Glu Glu Asp Leu
1 5 10 15
Pro Gly Glu Glu Asp Leu Pro Gly
20

<210> 15
<211> 13
<212> PRT
<213> HUMAN

<400> 15
Leu Glu Glu Gly Pro Glu Glu Asn Ser Ala Tyr Glu Gln
1 5 10

<210> 16
<211> 16
<212> PRT
<213> HUMAN

<400> 16
Met Arg Arg Gln His Arg Arg Gly Thr Lys Gly Gly Val Ser Tyr Arg
1 5 10 15

<210> 17
<211> 45
<212> DNA
<213> HUMAN

<400> 17
gtcgctagct ccatgggtca tatgcagagg ttgccccgga tgcag 45

<210> 18
<211> 43
<212> DNA
<213> HUMAN

<400> 18
gaagatctct tactcgagca ttctccaaga tccagcctct agg 43

<210> 19
<211> 10
<212> DNA
<213> HUMAN

<400> 19
ctccatctct 10

<210> 20
<211> 10
<212> DNA
<213> HUMAN

<400> 20
ccaccccat

10

<210> 21

<400> 21
000

<210> 22
<211> 8
<212> PRT
<213> HUMAN

<400> 22
Leu Glu His His His His His His
1 5

<210> 23
<211> 10
<212> DNA
<213> HUMAN

<220>
<221> misc_feature
<222> (1)..(10)

<400> 23
yyycayyyyy

10

<210> 24
<211> 10
<212> DNA
<213> HUMAN

<300>
<301> Locker and Buzard,
<303> DNA Sequencing and Mapping
<304> 1
<306> 3-11
<307> 1990

<400> 24
tgtgagactt

10

<210> 25
<211> 4

<212> PRT
<213> HUMAN

<220>

<221> SITE

<222> (1)..(4)

<223> sequence element defined by Suzuki, J. Mol. Biol., 207: 61-84 (1989) as motif frequently found in gene regulatory proteins.

<220>

<221> VARIANTS

<222> (3)..(4)

<223> variants in sequence element defined by Suzuki, J. Mol. Biol., 207: 61-84 (1989) as motif frequently found in gene regulatory proteins.

<400> 25

Ser Pro Xaa Xaa

1

<210> 26

<211> 4

<212> PRT

<213> HUMAN

<220>

<221> SITE

<222> (1)..(4)

<223> sequence element defined by Suzuki, J. Mol. Biol., 207: 61-84 (1989) as a motif frequently found in gene regulatory proteins.

<220>

<221> VARIANTS

<222> (3)..(4)

<223> variants in sequence element defined by Suzuki, J. Mol. Biol., 207: 61-84 (1989) as a motif frequently found in gene regulatory proteins.

<400> 26

Thr Pro Xaa Xaa

1

<210> 27

<211> 540

<212> DNA

<213> HUMAN

<220>

<221> promoter

<222> (1)..(540)

<400> 27

cttgcttttc attcaagctc aagtttgtct ccacataacc cattacttaa ctcaccctcg 60
ggctccccta gcagcctgcc ctacctcttt acctgcttcc tggaggagtc agggatgtat 120
acatgagctg ctttccctct cagccagagg acatgggggg cccagctcc cctgcctttc 180
cccttctgtg cctggagctg ggaagcaggc cagggttagc tgaggctggc tggcaagcag 240

```

ctgggtggtg ccagggagag cctgcatagt gccaggtggt gccttggggt ccaagctagt 300
ccatggcccc gataaccttc tgctgtgca cacacctgcc cctcactcca ccccatcct 360
agctttggtg tgggggagag ggcacagggc cagacaaacc tgtgagactt tggctccatc 420
tctgcaaaag ggcgctctgt gagtcagcct gctcccctcc aggcttgctc ctccccacc 480
cagctctcgt ttccaatgca cgtacagccc gtacacaccg tgtgctggga caccacacag 540

```

```

<210> 28
<211> 445
<212> DNA
<213> HUMAN

```

```

<220>
<221> exon
<222> (1)
<223> 1st MN exon

```

```

<400> 28
gcccgtagac accgtgtgct gggacacccc acagtcagcc gcatggctcc cctgtgcccc 60
agccccctggc tccctctgtt gatcccgccc cctgtctcag gcctcactgt gcaactgctg 120
ctgtcactgc tgcttctggt gcctgtccat cccagaggt tggcccgat gcaggaggat 180
tcccccttgg gaggaggctc ttctggggaa gatgaccac tgggcgagga ggatctgccc 240
agtgaagagg attcaccag agaggaggat ccaccggag aggaggatct acctggagag 300
gaggatctac ctggagagga ggatctacct gaagttaagc ctaaatacaga agaagagggc 360
tccctgaagt tagaggatct acctactgtt gaggctcctg gagatcctca agaaccacag 420
aataatgccc acagggacaa agaag                                     445

```

```

<210> 29
<211> 30
<212> DNA
<213> HUMAN

```

```

<220>
<221> exon
<222> (1)
<223> 2nd MN exon

```

```

<400> 29
gggatgacca ggtcattgg cgctatggag                                     30

```

```

<210> 30
<211> 171
<212> DNA
<213> HUMAN

```

```

<220>
<221> exon
<222> (1)
<223> 3rd MN exon

```

```

<400> 30
gggacccgcc ctggccccgg gtgtccccag cctgcgcggg ccgcttcag tccccgggtg 60
atatccgccc ccagctcgcc gccttctgcc cggccctgcg cccctggaa ctctgggct 120
tccagctccc gccgctccca gaactgcgcc tgcgcaacaa tggccacagt g                                     171

```


<210> 31
<211> 143
<212> DNA
<213> HUMAN

<220>
<221> exon
<222> (1)
<223> 4th MN exon

<400> 31
tgcaactgac cctgcctcct gggctagaga tggctctggg tcccgggagg gagtaccggg 60
ctctgcagct gcatctgcac tggggggctg caggctgtcc gggctcggag cacactgtgg 120
aaggccaccg tttccctgcc gag 143

<210> 32
<211> 93
<212> DNA
<213> HUMAN

<220>
<221> exon
<222> (1)
<223> 5th MN exon

<400> 32
atccacgtgg ttcacctcag caccgccttt gccagagttg acgaggcctt ggggagcccc 60
ggaggccttg ccgtgttggc cgcctttctg gag 93

<210> 33
<211> 67
<212> DNA
<213> HUMAN

<220>
<221> exon
<222> (1)
<223> 6th MN exon

<400> 33
gagggcccg aagaaaacag tgcctatgag cagttgctgt ctcgcttgga agaaatcgct 60
gaggaag 67

<210> 34
<211> 158
<212> DNA
<213> HUMAN

<220>
<221> exon
<222> (1)
<223> 7th MN exon

<400> 34
gctcagagac tcaggtccca ggactggaca tatctgcact cctgccctct gacttcagcc 60
gctacttcca atatgagggg tctctgacta caccgcctg tgcccagggt gtcacttgga 120
ctgtgtttaa ccagacagtg atgctgagtg ctaagcag 158

<210> 35
<211> 145
<212> DNA
<213> HUMAN

<220>
<221> exon
<222> (1)
<223> 8th MN exon

<400> 35
ctccacaccc tctctgacac cctgtgggga cctgggtgact ctcggctaca gctgaacttc 60
cgagcgacgc agcctttgaa tgggcgagtg attgaggcct ccttccttgc tggagtggac 120
agcagtcctc gggctgctga gccag 145

<210> 36
<211> 27
<212> DNA
<213> HUMAN

<220>
<221> exon
<222> (1)
<223> 9th MN exon

<400> 36
tccagctgaa ttcttgcttg gctgctg 27

<210> 37
<211> 82
<212> DNA
<213> HUMAN

<220>
<221> exon
<222> (1)
<223> 10th MN exon

<400> 37
gtgacatcct agccctgggt tttggcctcc tttttgctgt caccagcgtc gcgttccttg 60
tgcagatgag aaggcagcac ag 82

<210> 38
<211> 191
<212> DNA
<213> HUMAN

<220>

<221> exon
 <222> (1)
 <223> 11th MN exon

<400> 38
 aaggggaacc aaagggggtg tgagctaccg cccagcagag gtagccgaga ctggagccta 60
 gaggctggat cttggagaat gtgagaagcc agccagaggc atctgagggg gagccggtaa 120
 ctgtcctgtc ctgtcatta tgccacttcc ttttaactgc caagaaattt tttaaaataa 180
 atatttataa t 191

<210> 39
 <211> 1174
 <212> DNA
 <213> HUMAN

<220>
 <221> intron
 <222> (1)..(1174)
 <223> 1st MN intron

<400> 39
 gtaagtggct atcaatctcc aaatccaggt tccaggaggt tcatgactcc cctcccatac 60
 cccagcctag gctctgttca ctcagggagag gaggggagac tgtactcccc acagaagccc 120
 ttccagaggt cccataccaa tatccccatc cccactctcg gaggtagaaa gggacagatg 180
 tggagagaaa ataaaaaggg tgcaaaagga gagaggtgag ctggatgaga tgggagagaa 240
 gggggaggct ggagaagaga aagggatgag aactgcagat gagagaaaaa atgtgcagac 300
 agaggaaaaa aataggtgga gaaggagagt cagagagttt gaggggaaga gaaaaggaaa 360
 gcttgggagg tgaagtgggt accagagaca agcaagaaga gctggtagaa gtcattctcat 420
 cttaggctac aatgaggaat tgagacctag gaagaaggga cacagcaggt agagaaacgt 480
 ggcttcttga ctcccaagcc aggaatttgg ggaaaggggt tggagaccat acaaggcaga 540
 gggatgagtg gggagaagaa agaagggaga aaggaaagat ggtgtactca ctcatttggg 600
 actcaggact gaagtgccca ctcacttttt tttttttttt ttttgagaca aacttttact 660
 tttgttgccc aggctggagt gcaatggcgc gatctcggct cactgcaacc tccacctccc 720
 gggttcaagt gattctcctg cctcagcctc tagccaagta gctgcgatta caggcatgcg 780
 ccaccacgcc cggctaattt ttgtattttt agtagagacg gggtttcgcc atgttggtca 840
 ggctgggtctc gaactcctga tctcaggtga tccaaccacc ctggcctccc aaagtgcctg 900
 gattataggg gtgagccaca gcgcctggcc tgaagcagcc actcactttt acagacccta 960
 agacaatgat tgcaagctgg taggattgct gtttggccca cccagctgcg gtgttgagtt 1020
 tgggtgcggc ctctgtgct ttgcacctgg cccgcttaag gcatttgta cccgtaatgc 1080
 tcctgtaagg catctgcgtt tgtgacatcg ttttggtcgc caggaaggga ttggggctct 1140
 aagcttgagc ggttcatcct tttcatttat acag 1174

<210> 40
 <211> 193
 <212> DNA
 <213> HUMAN

<220>
 <221> intron
 <222> (1)..(193)
 <223> 2nd MN intron

<400> 40
 gtgagacacc caccgctgc acagacccaa tctgggaacc cagctctgtg gatctcccct 60
 acagccgtcc ctgaacactg gtcccgggcy tcccaccgcg cgcaccacgt cccaccccct 120

caccttttct acccgggttc cctaagttcc tgacctaggc gtcagacttc ctactatac 180
tctcccaccc cag 193

<210> 41
<211> 131
<212> DNA
<213> HUMAN

<220>
<221> intron
<222> (1)..(131)
<223> 3rd MN intron

<400> 41
gtgagggggt ctccccgccg agacttgggg atggggcggg gcgcagggaa gggaaccgtc 60
gcgcagtgcc tgccccgggg ttgggctggc cctaccgggc ggggccggct cacttgccct 120
tcctacgca g 131

<210> 42
<211> 89
<212> DNA
<213> HUMAN

<220>
<221> intron
<222> (1)..(89)
<223> 4th MN intron

<400> 42
gtgagcgcgg actggccgag aaggggcaaa ggagcggggc ggacgggggc cagagacgtg 60
gccctctcct accctcgtgt ccttttcag 89

<210> 43
<211> 1400
<212> DNA
<213> HUMAN

<220>
<221> intron
<222> (1)..(1400)
<223> 5th MN intron

<400> 43
gtaccagatc ctggacaccc cctactcccc gctttcccat cccatgctcc tcccggactc 60
tatcgtggag ccagagaccc catcccagca agctcactca ggcccctggc tgacaaactc 120
attcacgcac tgtttggtca tttaacaccc actgtgaacc aggcaccagc cccaacaag 180
gattctgaag ctgtaggctc ttgcctctaa ggagcccaca gccagtgggg gaggctgaca 240
tgacagacac ataggaagga catagtaaag atgggtggtc cagaggaggt gacacttaaa 300
gccttcactg gtagaaaaga aaaggaggtg ttcattgcag aggaaacaga atgtgcaaag 360
actcagaata tggcctatct aggggaatggc tacatacacc atgattagag gaggcccagt 420
aaagggaagg gatggtgaga tgccctgctag gttcactcac tcacttttat ttattttatt 480
atTTTTTTga cagtctctct gtcgcccagg ctggagtgcg gtggtgtgat cttgggtcac 540
tgcaacttcc gcctcccggt ttcaagggat tctcctgcct cagcttcctg agtagctggg 600
gttacagggt tgtgccacca tgcccagcta atTTTTTTTT gtatttttag tagacagggt 660

ttcaccatgt	tggtcaggct	ggtctcaaac	tcctggcctc	aagtgatccg	cctgactcag	720
cctaccaaag	tgctgattac	aagtgtgagc	caccgtgccc	agccacactc	actgattctt	780
taatgccagc	cacacagcac	aaagttcaga	gaaatgcctc	catcatagca	tgtcaatatg	840
ttcatactct	taggttcatg	atgttcttaa	cattaggttc	ataagcaaaa	taagaaaaaa	900
gaataataaa	taaaagaagt	ggcatgtcag	gacctcacct	gaaaagccaa	acacagaatc	960
atgaagggtga	atgcagaggt	gacaccaaca	caaagggtga	tatatggttt	cctgtgggga	1020
gtatgtacgg	aggcagcagt	gagtgagact	gcaaacgtca	gaagggcacg	ggtcactgag	1080
agcctagtat	cctagtaaag	tgggctctct	ccctctctct	ccagcttgtc	attgaaaacc	1140
agtccaccaa	gcttgttggt	tcgcacagca	agagtacata	gagtttgaaa	taatacatag	1200
gattttaaga	gggagacact	gtctctaaaa	aaaaaaacaa	cagcaacaac	aaaaagcaac	1260
aaccattaca	attttatggt	ccctcagcat	tctcagagct	gaggaatggg	agaggactat	1320
gggaaccccc	ttcatgttcc	ggccttcagc	catggccctg	gatacatgca	ctcatctgtc	1380
ttacaatgtc	attccccag					1400

<210> 44
 <211> 1334
 <212> DNA
 <213> HUMAN

<220>
 <221> intron
 <222> (1)..(1334)
 <223> 6th MN intron

<400> 44		
gtcagtttgt	tggtctggcc actaatctct gtggcctagt tcataaagaa tcaccctttg 60	
gagcttcagg	tctgaggctg gagatgggct ccctccagtg caggagggat tgaagcatga 120	
gccagcgctc	atcttgataa taacctatgaa gctgacagac acagttaccc gcaaaccggc 180	
gcctacagat	tgaaaaccaa gcaaaaaccg ccgggcacgg tggtcacgc ctgtaatccc 240	
agcacttttg	gaggccaagg cagggtggat acgagggtcaa gagatcaaga ccatcctggc 300	
caacatggtg	aaaccccatc tctactaaaa atacgaaaaa atagccaggc gtggtggcgg 360	
gtgcctgtaa	tcccagctac tcgggaggct gaggcaggag aatggcatga acccgggagg 420	
cagaagttgc	agtgagccga gatcgtgcc a ctgcactcca gcctgggcaa cagagcgaga 480	
ctcttgtctc	aaaaaaaaaa aaaaaaaaga aaaccaagca aaaacaaaa tgagacaaaa 540	
aaaacaagac	caaaaaatgg tgtttggaaa ttgtcaaggt caagtctgga gagctaaact 600	
ttttctgaga	actgtttatc ttttaataagc atcaaatatt ttaactttgt aaatactttt 660	
gttggaatc	gttctcttct tagtcactct tgggtcattt taaatctcac ttactctact 720	
agacctttta	ggtttctgct agactaggta gaactctgcc tttgcatttc ttgtgtctgt 780	
tttgtatagt	tatcaatatt catatttatt tacaagttat tcagatcatt ttttcttttc 840	
tttttttttt	tttttttttt ttttacatct ttagtagaga cagggtttca ccatattggc 900	
caggctgctc	tcaaactcct gaccttgtga tccaccagcc tcggcctccc aaagtgtctg 960	
gattcatttt	ttctttttta tttgtctctg gcttaaactt gtggcccagc actttatgat 1020	
ggtacacaga	gttaagagtg tagactcaga cggctcttct tctttccttc tcttccttcc 1080	
tcccttccct	cccaccttcc cttctctcct tcctttcttt ctctctctct tgcttctca 1140	
ggcctcttcc	agttgctcca aagccctgta cttttttttg agttaacgtc ttatgggaag 1200	
ggcctgcact	tagtgaagaa gtggtctcag agttgagtta ccttggtctc tgggagggtga 1260	
aactgtatcc	ctataccctg aagctttaag ggggtgcaat gtagatgaga ccccaacata 1320	
gacctcttcc	acag	1334

<210> 45
 <211> 512
 <212> DNA
 <213> HUMAN

<220>

<221> intron
 <222> (1)..(512)
 <223> 7th MN intron

<400> 45
 gtgggcctgg ggtgtgtgtg gacacagtgg gtgcggggga aagaggatgt aagatgagat 60
 gagaaacagg agaagaaaga aatcaaggct gggctctgtg gcttacgcct ataatcccac 120
 cacgttggga ggctgaggtg ggagaatggt ttgagcccag gagttcaaga caaggcgggg 180
 caacatagtg tgaccccatc tctacaaaa aaacccaac aaaacaaaa atagccgggc 240
 atggtggtat gcggcctagt cccagctact caaggaggct gaggtgggaa gatcgcttga 300
 ttccaggagt ttgagactgc agtgagctat gatcccacca ctgcctacca tcttttaggat 360
 acattttattt atttataaaa gaaatcaaga ggctggatgg ggaatacagg agctggaggg 420
 tggagccctg aggtgctggt tgtgagctgg cctgggaccc ttgtttcctg tcatgccatg 480
 aaccaccca cactgtccac tgacctccct ag 512

<210> 46
 <211> 114
 <212> DNA
 <213> HUMAN

<220>
 <221> intron
 <222> (1)..(114)
 <223> 8th MN intron

<400> 46
 gtacagcttt gtctggtttc cccccagcca gtagtccctt atcctcccat gtgtgtgcc 60
 gtgtctgtca ttggtggtca cagcccgct ctcacatctc ctttttctct ccag 114

<210> 47
 <211> 617
 <212> DNA
 <213> HUMAN

<220>
 <221> intron
 <222> (1)..(617)
 <223> 9th MN intron

<400> 47
 gtgagtctgc cctcctctt ggtcctgatg ccaggagact cctcagcacc attcagcccc 60
 agggctgctc aggaccgct ctgctccctc tccttttctg cagaacagac cccaaccca 120
 atattagaga ggcagatcat ggtggggatt ccccatgtt cccagaggc taattgatta 180
 gaatgaagct tgagaaatct cccagcatcc ctctcgcaa agaatcccc cccctttttt 240
 taaagatagg gtctcactct gtttgcccca ggctggggtg ttgtggcacg atcatagctc 300
 actgcagcct cgaactccta ggctcaggca atcctttcac cttagcttct caaagcactg 360
 ggactgtagg catgagccac tgtgcctggc ccaaacggc ccttttactt ggcttttagg 420
 aagcaaaaac ggtgcttata ttaccccttc tctgtatat accctcatcc cttggctggc 480
 ctcttctgga gactgaggca ctatggggct gcctgagaac tcggggcagg ggtggtggag 540
 tgcactgagg caggtgttga ggaactctgc agaccctct tccttccaa agcagccctc 600
 tctgctctcc atcgag 617

<210> 48
 <211> 130

<212> DNA
<213> HUMAN

<220>
<221> intron
<222> (1)..(130)
<223> 10th MN intron

<400> 48
gtattacact gaccctttct tcaggcacia gcttcccca cccttgtgga gtcacttcat 60
gcaaagcgca tgcaaatgag ctgctcctgg gccagtttcc tgattagcct ttccctgttgt 120
gtacacacag 130

<210> 49
<211> 1401
<212> DNA
<213> HUMAN

<400> 49
caaactttca cttttgttgc ccaggctgga gtgcaatggc gcgatctcgg ctactgcaa 60
cctccacctc ccgggttcaa gtgattctcc tgcctcagcc tctagccaag tagctgcgat 120
tacaggcatg cgccaccacg cccggctaatt ttttgtattt ttagtagaga cgggggtttcg 180
ccatgttggt caggctgggc tcgaactcct gatctcaggt gatccaacca ccctggcctc 240
ccaaagtgcg gggattatag gcgtgagcca cagcgcctgg cctgaagcag ccactcactt 300
ttacagacct taagacaatg attgcaagct ggtaggattg ctgtttggcc caccagctg 360
cgggtgttgag tttgggtgag gtctcctgtg ctttgcacct ggcccgtta aggcatttgt 420
taccgtaat gctcctgtaa ggcactctgcg tttgtgacat cgttttgggc gccaggaagg 480
gattggggct ctaagcttga gcggttcac cttttcattt atacagggga tgaccagagt 540
cattggcgct atggaggtga gacaccacc cgctgcacag acccaatctg ggaaccagc 600
tctgtggatc tcccctacag ccgtccctga acactgggtc cgggcgtccc accgcgcgc 660
caccgtccca ccccctacc ttttctaccc ggggtcccta agttcctgac ctaggcgtca 720
gacttcctca ctatactctc ccaccccagg cgaccgcgcc tggccccggg tgtcccagc 780
ctgcgcgggc cgcttccagt ccccggtgga tatccgcccc cagctcgccg ccttctgccc 840
ggccctgcgc cccctggaac tccctgggctt ccagctcccg ccgctcccag aactgcgcct 900
gcgcaacaat ggccacagtg gtgagggggt ctccccgcgc agacttgggg atggggcggg 960
gcgcagggaa gggaaccgtc gcgcagtgcc tgcccggggg ttgggctggc cctaccgggc 1020
ggggccggct cacttgctc tccctacgca gtgcaactga ccctgcctcc tgggctagag 1080
atggctctgg gtcccgggcg ggagtaccgg gctctgcagc tgcacttgca ctggggggct 1140
gcaggtcgct cgggctcgga gcacactgtg gaaggccacc gtttccctgc cgaggtgagc 1200
gcggactggc cgagaagggg caaaggagcg gggcgagcgg gggccagaga cgtggccctc 1260
tcctaccctc gtgtcctttt cagatccacg tggttcacct cagcaccgcc tttgccagag 1320
ttgacgaggg cttggggcgc ccgggaggcc tggccgtgtt ggccgccttt ctggaggtac 1380
cagatcctgg acacccccta c 1401

<210> 50
<211> 59
<212> PRT
<213> HUMAN

<400> 50
Ser Ser Gly Glu Asp Asp Pro Leu Gly Glu Glu Asp Leu Pro Ser Glu
1 5 10 15
Glu Asp Ser Pro Arg Glu Glu Asp Pro Pro Gly Glu Glu Asp Leu Pro
20 25 30

Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro Glu Val Lys Pro
 35 40 45

Lys Ser Glu Glu Glu Gly Ser Leu Lys Leu Glu
 50 55

<210> 51
 <211> 257
 <212> PRT
 <213> HUMAN

<400> 51
 Gly Asp Asp Gln Ser His Trp Arg Tyr Gly Gly Asp Pro Pro Trp Pro
 1 5 10 15

Arg Val Ser Pro Ala Cys Ala Gly Arg Phe Gln Ser Pro Val Asp Ile
 20 25 30

Arg Pro Gln Leu Ala Ala Phe Cys Pro Ala Leu Arg Pro Leu Glu Leu
 35 40 45

Leu Gly Phe Gln Leu Pro Pro Leu Pro Glu Leu Arg Leu Arg Asn Asn
 50 55 60

Gly His Ser Val Gln Leu Thr Leu Pro Pro Gly Leu Glu Met Ala Leu
 65 70 75 80

Gly Pro Gly Arg Glu Tyr Arg Ala Leu Gln Leu His Leu His Trp Gly
 85 90 95

Ala Ala Gly Arg Pro Gly Ser Glu His Thr Val Glu Gly His Arg Phe
 100 105 110

Pro Ala Glu Ile His Val Val His Leu Ser Thr Ala Phe Ala Arg Val
 115 120 125

Asp Glu Ala Leu Gly Arg Pro Gly Gly Leu Ala Val Leu Ala Ala Phe
 130 135 140

Leu Glu Glu Gly Pro Glu Glu Asn Ser Ala Tyr Glu Gln Leu Leu Ser
 145 150 155 160

Arg Leu Glu Glu Ile Ala Glu Glu Gly Ser Glu Thr Gln Val Pro Gly
 165 170 175

Leu Asp Ile Ser Ala Leu Leu Pro Ser Asp Phe Ser Arg Tyr Phe Gln
 180 185 190

Tyr Glu Gly Ser Leu Thr Thr Pro Pro Cys Ala Gln Gly Val Ile Trp
 195 200 205

Thr Val Phe Asn Gln Thr Val Met Leu Ser Ala Lys Gln Leu His Thr
 210 215 220

Leu Ser Asp Thr Leu Trp Gly Pro Gly Asp Ser Arg Leu Gln Leu Asn

<213> HUMAN

<400> 55

```
cauggccccc auaaccuucu gccugugcac acaccugccc cucacuccac ccccauccua 60
gcuuugguau gggggagagg gcacagggcc agacaaaccu gugagacuuu ggcuccaucu 120
cugcaaaagg ggcucucugug agucagccug cuccccucca ggcugucucc uccccacccc 180
agcucucguu uccaauGCac guacagcccG uacacaccgu gugcugggac accccacagu 240
cagccgcaug gcuccccugu gcccagccc cuggcucccu cuguugauc cggccccugc 300
uccagggcuc acugugcaac ugcugcuguc acugcugcuu cuggugccug uccaucacca 360
gagguugccc cggauGCagg aggauucGcc cuugggagga ggcucuucug gggaagauga 420
cccacugggc gaggaggauC ugcccaguga agaggauuca cccagagagg 470
```

<210> 56

<211> 292

<212> DNA

<213> HUMAN

<400> 56

```
gttttttttga gacggagtct tgcattctgtc atgccaggcc tggagttagca gtggtgccat 60
ctcggctcac tgcaagctcc acctcccgag ttcacgccat tttcctgcct cagcctcccc 120
agtagctggg actacaggcg cccgccacca tgcccggtta attttttgta tttttggtag 180
agacgggggtt tcaccgtgtt agccagaatg gtctcgatct cctgacttcg tgatccaccc 240
gcctcggcct cccaaagttc tgggattaca ggtgtgagcc accgcacctg gc 292
```

<210> 57

<211> 262

<212> DNA

<213> HUMAN

<400> 57

```
tttctttttt gagacagggt cttgctctgt caccagggcc agagtgcaat ggtacagtct 60
cagctcactg cagcctcaac cgctcggct caaacatca tccatttca gcctcctgag 120
tagctgggac tacaggcaca tgccattaca cctggcta atttttgtat ttctagtaga 180
gacagggttt ggccatgttg cccgggctgg tctcgaactc ctggactcaa gcaatccacc 240
cacctcagcc tcccaaatg ag 262
```

<210> 58

<211> 2501

<212> DNA

<213> HUMAN

<220>

<221> misc_feature

<222> (1)..(2501)

<223> region 5' to transcription initiation site as determined by RNase protection assay (nucleotide 3507 of Figures 2A-2F and of SEQ ID NO: 5), corresponding to region of SEQ ID NO: 5 and Figures 2A-2F from nucleotide (7) to nucleotide (2507), in which region some regulatory elements are probably situated.

<220>

<221> unsure what base is at position 1968

<222> (1968)

<223> unsure of base at position 1968, which is the same unknown base as that at position 1974 of SEQ ID NO. 5, i.e., the full-length MN genomic sequence, and of that unknown at position 1968 of SEQ ID NO: 90, and unknown at position 647 of SEQ ID NO: 110. That unknown base is in the 5' region flanking the transcription initiation site (3507) as determined by RNase protection assay.

<400> 58

```

tgttgactcg tgaccttacc cccaaccctg tgctctctga aacatgagct gtgtccactc 60
aggggttaa at ggattaaggg cgggtgcaaga tgtgctttgt taaacagatg cttgaaggca 120
gcatgctcgt taagagtcac caccaatccc taatctcaag taatcaggga cacaacact 180
gcggaaggcc gcagggtcct ctgcctagga aaaccagaga cctttgttca cttgtttatc 240
tgaccttccc tccactattg tccatgaccc tgccaaatcc cctctgtga gaaacaccca 300
agaattatca ataaaaaaat aaatttataa aaaaaatata aaaaaaaaaa aaaaaaaaaa 360
aaaagactta cgaatagtta ttgataaatg aatagctatt ggtaaagcca agtaaagtat 420
catattcaaa accagacggc catcatcaca gctcaagtct acctgatttg atctctttat 480
cattgtcatt ctttggattc actagattag tcatcatcct caaaattctc ccccaagttc 540
taattacgtt ccaaacattt aggggttaca tgaagcttga acctactacc ttctttgctt 600
ttgagccatg agttgttaga atgatgagtt tacaccttac atgctgggga ttaattttaa 660
ctttacctct aagtcagttg ggtagccttt ggcttatttt tgtagctaat tttgtagtta 720
atggatgcac tgtgaatctt gctatgatag ttttctcca cactttgcca ctaggggtag 780
gtaggtactc agttttcagt aattgcttac ctaagaccct aagccctatt tctcttgtag 840
tggcctttat ctgtaatatg ggcataattt atacaatata atttttggag tttttttgtt 900
tgtttgtttg tttgtttttt tgagacggag tcttgcatct gtcatgcca ggctggagta 960
gcagtggtgc catctcggct cactgcaagc tccacctccc gagttcacgc cattttctctg 1020
cctcagcctc ccgagtagct gggactacag gcgccgccca ccatgcccgg ctaatttttt 1080
gtattttttg tagagacggg gtttcaccgt gtttagccaga atggtctcga tctcctgact 1140
tcgtgatcca ccgcctcgg cctcccaaag ttctgggatt acagggtgtg gccaccgcac 1200
ctggccaatt ttttgagctt tttaaagtaa aaatatgtct tgtaagctgg taactatggg 1260
acatttctct ttattaatgt ggtgctgacg gtcatatagg ttcttttgag tttggcatgc 1320
atatgctact ttttgcagtc ctttcattac attttctct cttcatttga agagcatgtt 1380
atatctttta gcttcacttg gcttaaaagg ttctctcatt agcctaacac agtgtcattg 1440
ttggtaccac ttggatcata agtggaaaaa cagtcaagaa attgcacagt aatacttggt 1500
tgtaagaggg atgattcagg tgaatctgac actaagaaac tccctacct gaggtctgag 1560
attcctctga cattgctgta tataggcttt tcttttgaca gcctgtgact gcggactatt 1620
tttcttaagc aagatatgct aaagttttgt gagecttttt ccagagagag gtctcatatc 1680
tgcacaaagt gagaacatat aatgtctgca tgtttccata tttcaggaat gtttgcttgt 1740
gttttatgct tttatataga cagggaact tgttctcag tgacccaaaa gaggtgggaa 1800
ttgttatttg atatcatcat tggcccacgc tttctgacct tggaaacaat taagggttca 1860
taatctcaat tctgtcagaa ttggtacaag aaatagctgc tatgtttctt gacattccac 1920
ttggtaggaa ataagaatgt gaaactcttc agttggtgtg tgtccctngt ttttttgcaa 1980
tttcttctct actgtgttaa aaaaaagtat gatcttgctc tgagaggtga ggcattctta 2040
atcatgatct ttaaagatca ataataaat cctttcaagg attatgtctt tattataata 2100
aagataatth gtctttaaca gaatcaataa tataatccct taaaggatta tatctttgct 2160
gggcgcagtg gctcacacct gtaatcccag cactttgggt ggccaagggt gaaggatcaa 2220
atttgcttac ttctatatta tcttctaaag cagaattcat ctctcttccc tcaatatgat 2280
gatattgaca gggtttgccc tcaactacta gattgtgagc tctgtctcag ggcaggtagc 2340
gttttttggt tttgtttttg tttttctttt ttgagacagg gtcttgctct gtcacccagg 2400
ccagagtgca atggtacagt ctcagctcac tgcagcctca accgcctcgg ctcaaaccat 2460
catcccattt cagcctcctg agtagctggg actacaggca c 2501

```

<210> 59

<211> 292

<212> DNA

<213> HUMAN

<220>

<221> misc_feature

<222> (1)

<400> 59

```
ttttttttgag acggagtctt gcatctgtca tgcccaggct ggagtagcag tgggtgccatc 60
tcggctcact gcaagctcca cctcccagat tcacgccatt ttccctgcctc agcctcccga 120
gtagctggga ctacaggcgc ccgccaccat gcccggttaa ttttttgtat ttttggtaga 180
gacgggggtt caccgtgtta gccagaatgg tctcgatctc ctgacttcgt gatccacccg 240
cctcggcctc ccaaagttct gggattacag gtgtgagcca ccgcacctgg cc 292
```

<210> 60

<211> 262

<212> DNA

<213> HUMAN

<400> 60

```
ttcttttttg agacagggtc ttgctctgtc acccaggcca gagtgcaatg gtacagtctc 60
agctcactgc agcctcaacc gcctcggctc aaaccatcat cccatttcag cctcctgagt 120
agctgggact acaggcacat gccattacac ctggctaatt tttttgtatt tctagtagag 180
acaggggttg gccatgttgc ccgggctggt ctcgaactcc tggactcaag caatccaccc 240
acctcagcct cccaaaatga gg 262
```

<210> 61

<211> 294

<212> DNA

<213> HUMAN

<400> 61

```
tttttttttg agacaaactt tcacttttgt tgcccaggct ggagtgcaat ggcgcgatct 60
cggctcactg caacctccac ctcccgggtt caagtgatc tcctgcctca gcctctagcc 120
aagtagctgc gattacaggc atgcgccacc acgcccggct aatttttgta tttttagtag 180
agacgggggt tcgccatgtt ggtcaggctg gtctcgaact cctgatctca ggtgatccaa 240
ccacctgggc ctcccaaagt gctgggatta taggcgtgag ccacagcgcc tggc 294
```

<210> 62

<211> 276

<212> DNA

<213> HUMAN

<400> 62

```
tgacagtctc tctgtcgccc aggctggagt gcagtgggtg gatcttgggt cactgcaact 60
tcgcctcccc gggttcaagg gattctcctg cctcagcttc ctgagtagct ggggttacag 120
gtgtgtgcca ccatgcccag ctaatttttt tttgtatttt tagtagacag ggtttcacca 180
tgttggtcag gctgggtctc aactcctggc ctcaagtgat ccgcctgact cagcctacca 240
aagtgtgat tacaagtgtg agccaccgtg cccagc 276
```

<210> 63

<211> 289

<212> DNA

<213> HUMAN

<400> 63

```
cgccgggcac ggtggctcac gcctgtaatc ccagcacttt gggaggccaa ggcaggtgga 60
```

tcacgaggtc aagagatcaa gaccatcctg gccaacatgg tgaaacccca tctctactaa 120
aaatacgaaa aaatagccag gcgtggtggc ggggtgcctgt aatcccagct actcgggagg 180
ctgaggcagg agaatggcat gaaccggga ggcagaagtt gcagtgcagc gagatcgtgc 240
cactgcactc cagcctgggc aacagagcga gactcttgct tcaaaaaa 289

<210> 64
<211> 298
<212> DNA
<213> HUMAN

<400> 64
aggctgggct ctgtggctta cgcctataat cccaccacgt tgggaggctg aggtgggaga 60
atggtttgag ccaggagtt caagacaagg cggggcaaca tagtgtagc ccatctctac 120
caaaaaaacc ccaacaaaac caaaaatagc cgggcatggt ggtatgcggc ctagtcccag 180
ctactcaagg aggtgaggt gggaagatcg cttgattcca ggagtttgag actgcagtga 240
gctatgatcc caccactgcc taccatcttt aggatacatt tatttattta taaaagaa 298

<210> 65
<211> 105
<212> DNA
<213> HUMAN

<400> 65
ttttttacat ctttagtaga gacagggttt caccatattg gccaggctgc tctcaaactc 60
ctgaccttgt gatccaccag cctcggcctc ccaaagtgt gggat 105

<210> 66
<211> 83
<212> DNA
<213> HUMAN

<400> 66
cctcgaactc ctaggctcag gcaatccttt caccttagct tctcaaagca ctgggactgt 60
aggcatgagc cactgtgcct ggc 83

<210> 67
<211> 11
<212> DNA
<213> HUMAN

<400> 67
agaaggtgag t 11

<210> 68
<211> 11
<212> DNA
<213> HUMAN

<400> 68
tggaggtgag a 11

<210> 69
<211> 11
<212> DNA
<213> HUMAN

<400> 69
cagtcgtgag g

11

<210> 70
<211> 11
<212> DNA
<213> HUMAN

<400> 70
ccgaggtgag c

11

<210> 71
<211> 11
<212> DNA
<213> HUMAN

<400> 71
tggaggtacc a

11

<210> 72
<211> 11
<212> DNA
<213> HUMAN

<400> 72
ggaaggtcag t

11

<210> 73
<211> 11
<212> DNA
<213> HUMAN

<400> 73
agcaggtggg c

11

<210> 74
<211> 11
<212> DNA
<213> HUMAN

<400> 74
gccaggtaca g

11

<210> 75
<211> 11
<212> DNA

<213> HUMAN

<400> 75

tgctggtgag t

11

<210> 76

<211> 11

<212> DNA

<213> HUMAN

<400> 76

cacaggtatt a

11

<210> 77

<211> 11

<212> DNA

<213> HUMAN

<400> 77

atacagggga t

11

<210> 78

<211> 11

<212> DNA

<213> HUMAN

<400> 78

ccccaggcga c

11

<210> 79

<211> 11

<212> DNA

<213> HUMAN

<400> 79

acgcagtgca a

11

<210> 80

<211> 11

<212> DNA

<213> HUMAN

<400> 80

tttcagatcc a

11

<210> 81

<211> 11

<212> DNA

<213> HUMAN

<400> 81

ccccaggagg g 11

<210> 82
<211> 11
<212> DNA
<213> HUMAN

<400> 82
tcacaggctc a 11

<210> 83
<211> 11
<212> DNA
<213> HUMAN

<400> 83
ccctagctcc a 11

<210> 84
<211> 11
<212> DNA
<213> HUMAN

<400> 84
ctccagtcga g 11

<210> 85
<211> 12
<212> DNA
<213> HUMAN

<400> 85
tcgcaggtga ca 12

<210> 86
<211> 11
<212> DNA
<213> HUMAN

<400> 86
acacagaagg g 11

<210> 87
<211> 377
<212> PRT
<213> HUMAN

<400> 87
Gln Arg Leu Pro Arg Met Gln Glu Asp Ser Pro Leu Gly Gly Gly Ser
1 5 10 15

Ser Gly Glu Asp Asp Pro Leu Gly Glu Glu Asp Leu Pro Ser Glu Glu
 20 25 30
 Asp Ser Pro Arg Glu Glu Asp Pro Pro Gly Glu Glu Asp Leu Pro Gly
 35 40 45
 Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro Glu Val Lys Pro Lys
 50 55 60
 Ser Glu Glu Glu Gly Ser Leu Lys Leu Glu Asp Leu Pro Thr Val Glu
 65 70 75 80
 Ala Pro Gly Asp Pro Gln Glu Pro Gln Asn Asn Ala His Arg Asp Lys
 85 90 95
 Glu Gly Asp Asp Gln Ser His Trp Arg Tyr Gly Gly Asp Pro Pro Trp
 100 105 110
 Pro Arg Val Ser Pro Ala Cys Ala Gly Arg Phe Gln Ser Pro Val Asp
 115 120 125
 Ile Arg Pro Gln Leu Ala Ala Phe Cys Pro Ala Leu Arg Pro Leu Glu
 130 135 140
 Leu Leu Gly Phe Gln Leu Pro Pro Leu Pro Glu Leu Arg Leu Arg Asn
 145 150 155 160
 Asn Gly His Ser Val Gln Leu Thr Leu Pro Pro Gly Leu Glu Met Ala
 165 170 175
 Leu Gly Pro Gly Arg Glu Tyr Arg Ala Leu Gln Leu His Leu His Trp
 180 185 190
 Gly Ala Ala Gly Arg Pro Gly Ser Glu His Thr Val Glu Gly His Arg
 195 200 205
 Phe Pro Ala Glu Ile His Val Val His Leu Ser Thr Ala Phe Ala Arg
 210 215 220
 Val Asp Glu Ala Leu Gly Arg Pro Gly Gly Leu Ala Val Leu Ala Ala
 225 230 235 240
 Phe Leu Glu Glu Gly Pro Glu Glu Asn Ser Ala Tyr Glu Gln Leu Leu
 245 250 255
 Ser Arg Leu Glu Glu Ile Ala Glu Glu Gly Ser Glu Thr Gln Val Pro
 260 265 270
 Gly Leu Asp Ile Ser Ala Leu Leu Pro Ser Asp Phe Ser Arg Tyr Phe
 275 280 285
 Gln Tyr Glu Gly Ser Leu Thr Thr Pro Pro Cys Ala Gln Gly Val Ile
 290 295 300
 Trp Thr Val Phe Asn Gln Thr Val Met Leu Ser Ala Lys Gln Leu His
 305 310 315 320

Thr Leu Ser Asp Thr Leu Trp Gly Pro Gly Asp Ser Arg Leu Gln Leu
325 330 335

Asn Phe Arg Ala Thr Gln Pro Leu Asn Gly Arg Val Ile Glu Ala Ser
340 345 350

Phe Pro Ala Gly Val Asp Ser Ser Pro Arg Ala Ala Glu Pro Val Gln
355 360 365

Leu Asn Ser Cys Leu Ala Ala Gly Asp
370 375

<210> 88
<211> 34
<212> DNA
<213> HUMAN

<400> 88
tagacagatc tacgatggct cccctgtgcc ccag 34

<210> 89
<211> 34
<212> DNA
<213> HUMAN

<400> 89
attcctctag acagttaccg gctccccctc agat 34

<210> 90
<211> 3532
<212> DNA
<213> HUMAN

<220>
<221> misc_feature which includes the MN gene promoter
<222> (1)..(3532)
<223> region including the transcription initiation site (nucleotide 3507 of
SEQ ID NO: 5 and of Figures 2A-2F) as determined by RNase protection assay,
which region is inclusive of the MN gene promoter, and corresponds to
nucleotide 7 to nucleotide 3538 of SEQ ID NO: 5 and of Figures 2A-2F.

<220>
<221> unsure what base is at position 1968
<222> (1968)
<223> unsure of the base at position 1968, which is the same unknown base at
position 1974 of SEQ ID NO: 5 (the full-length MN genomic sequence), position
1968 of SEQ ID NO: 58 and position 647 of SEQ ID NO: 110. That unknown base
is in the region that includes the transcription initiation site (nucleotide
3507 of SEQ ID NO: 5 and of Figures 2A-2F) as determined by RNase protection
assay, which region is inclusive of the MN gene promoter.

<400> 90
tggtgactcg tgaccttacc cccaaccctg tgctctctga aacatgagct gtgtccactc 60
agggttaaat ggattaaggc cggtgcaaga tgtgctttgt taaacagatg cttgaaggca 120

gcattgctcgt	taagagtcatt	caccaatccc	taatctcaag	taatcaggga	cacaaacact	180
gcggaaggcc	gcagggtcct	ctgcctagga	aaaccagaga	cctttgttca	cttgtttatc	240
tgaccttccc	tccactattg	tccatgaccc	tgccaaatcc	ccctctgtga	gaaacaccca	300
agaattatca	ataaaaaaat	aaatttaaaa	aaaaaataca	aaaaaaaaaa	aaaaaaaaaa	360
aaaagactta	cgaatagtta	ttgataaatg	aatagctatt	ggtaaagcca	agtaaattgat	420
catattcaaa	accagacggc	catcatcaca	gctcaagtct	acctgatttg	atctctttat	480
cattgtcatt	ctttggattc	actagattag	tcatcatcct	caaaattctc	ccccaaagttc	540
taattacgtt	ccaaacattt	aggggttaca	tgaagcttga	acctactacc	ttctttgctt	600
ttgagccatg	agttgttaga	atgatgagtt	tacaccttac	atgctgggga	ttaatTTaaa	660
ctttacctct	aagtcagttg	ggtagccttt	ggcttatttt	tgtagctaata	tttgtagtta	720
atggatgcac	tgtgaatctt	gctatgatag	ttttcctcca	cactttgcca	ctaggggtag	780
gtaggctactc	agttttcagt	aattgcttac	ctaagaccct	aagccctatt	tctcttgtag	840
tggcctttat	ctgtaatatg	ggcatattta	atacaatata	atTTTTggag	tttttttggt	900
tgtttggttg	tttgTTTTtt	tgagacggag	tcttgcatct	gtcatgcccc	ggctggagta	960
gcagtgggtgc	catctcggct	cactgcaagc	tccacctccc	gagttcacgc	cattttcctg	1020
cctcagcctc	ccgagtagct	gggactacag	gcgcccgcga	ccatgcccgg	ctaattTTTT	1080
gtattTTTTg	tagagacggg	gtttcacctg	gttagccaga	atggctctga	tctcctgact	1140
tcgtgatcca	ccgcctcgg	cctcccaaag	ttctgggatt	acaggtgtga	gccaccgcac	1200
ctggccaatt	ttttgagctc	tttaaagtaa	aaatatgtct	tgtaagctgg	taactatggg	1260
acatttctct	ttattaatgt	gggtgctgacg	gtcatatagg	ttcttttgag	tttggcatgc	1320
atatgctact	ttttgcagtc	ctttcattac	atTTTTctct	cttcatttga	agagcatggt	1380
atatctttta	gcttcacttg	gcttaaaagg	ttctctcatt	agcctaacac	agtgtcattg	1440
ttggtaccac	ttggatcata	agtggaaaaa	cagtcaagaa	attgcacagt	aatacttggt	1500
tgtaagaggg	atgattcagg	tgaatctgac	actaagaaac	tccctacct	gaggtctgag	1560
attcctctga	cattgctgta	tataggcttt	tcctttgaca	gcctgtgact	gcggactatt	1620
tttcttaagc	aagatatgct	aaagttttgt	gagccttttt	ccagagagag	gtctcatatc	1680
tgcatcaagt	gagaacatat	aatgtctgca	tgtttccata	tttcaggaat	gtttgcttgt	1740
gttttatgct	tttatataga	cagggaaact	tgttcctcag	tgacccaaaa	gaggtgggaa	1800
ttgttatttg	atatcatcat	tgccccacgc	tttctgacct	tggaacaat	taagggttca	1860
taatctcaat	tctgtcagaa	ttggtacaag	aaatagctgc	tatgtttctt	gacattccac	1920
ttggtaggaa	ataagaatgt	gaaactcttc	agttgggtgtg	tgccctngt	ttttttgcaa	1980
tttccttctt	actgtgttaa	aaaaaagtat	gatcttgctc	tgagaggtga	ggcattctta	2040
atcatgatct	ttaaagatca	ataatataat	cctttcaagg	attatgtctt	tattataata	2100
aagataattt	gtcttttaaca	gaatcaataa	tataatccct	taaaggatta	tatctttgct	2160
gggcgcagtg	gctcacacct	gtaatcccag	cactttgggt	ggccaagggtg	gaaggatcaa	2220
atttgcttac	ttctatatta	tcttctaaag	cagaattcat	ctctcttccc	tcaatatgat	2280
gatattgaca	gggtttgccc	tactcacta	gattgtgagc	tcctgtctcag	ggcaggtagc	2340
gttttttggt	tttgTTTTtt	tttttctttt	ttgagacagg	gtcttgctct	gtcaccacag	2400
ccagagtgca	atggtacagt	ctcagctcac	tgcagcctca	accgcctcgg	ctcaaaccat	2460
catcccatTT	cagcctcctg	agtagctggg	actacaggca	catgccatta	cacctggcta	2520
atTTTTttgt	atTTctagta	gagacagggt	ttggccatgt	tgcccgggct	ggctctgaac	2580
tcctggactc	aagcaatcca	cccacctcag	cctcccaaaa	tgagggaccg	tgtcttattc	2640
atTTccatgt	ccctagtcca	tagcccagtg	ctggacctat	ggtagtacta	aataaatatt	2700
tgttgaaatgc	aatagtaaat	agcattttcag	ggagcaagaa	ctagattaac	aaaggtggta	2760
aaaggttttg	agaaaaaaat	aatagtttaa	tttggttaga	gtatgaggga	gagtagtagg	2820
agacaagatg	gaaagggtctc	ttgggcaagg	ttttgaagga	agttggaagt	cagaagtaca	2880
caatgtgcat	atcgtggcag	gcagtgggga	gccaatgaag	gcttttgagc	aggagagtaa	2940
tgtgttgaaa	aataaatata	ggttaaaccct	atcagagccc	ctctgacaca	tacacttgct	3000
tttcattcaa	gctcaagttt	gtctcccaca	taccatttac	ttaaactcacc	ctcgggctcc	3060
cctagcagcc	tgccctacct	ctttacctgc	ttcctgggtg	agtcagggat	gtatacatga	3120
gctgctttcc	ctctcagcca	gaggacatgg	ggggccccag	ctccccctgcc	tttccccctc	3180
tgtgcctgga	ctctgggaagc	aggccagggt	tagctgaggc	tggtctggcaa	gcagctgggt	3240
gggtgccagg	agagcctgca	tagtgccagg	tggtgccttg	ggttccaagc	tagtccatgg	3300
ccccgataac	cttctgcttg	tgcacacacc	tgcctctcac	tccacccccca	tcctagcttt	3360
gggatggggg	agagggcaca	gggccagaca	aacctgtgag	actttggctc	catctctgca	3420
aaagggcgct	ctgtgagtca	gcctgctccc	ctccaggctt	gctcctcccc	caccagctc	3480
tcgtttccaa	tgcacgtaca	gcccgtagac	accgtgtgct	gggacacccc	ac	3532

<210> 91
 <211> 204
 <212> DNA
 <213> HUMAN

<400> 91
 cctgccccctc actccacccc catcctagct ttggtatggg ggagagggca cagggccaga 60
 caaacctgtg agactttggc tccatctctg caaaagggcg ctctgtgagt cagcctgctc 120
 ccctccaggc ttgctcctcc cccaccagc tctcgtttcc aatgcacgta cagcccgtac 180
 acaccgtgtg ctgggacacc ccac 204

<210> 92
 <211> 132
 <212> DNA
 <213> HUMAN

<400> 92
 ggatcctggt gactcgtgac cttaccccca accctgtgct ctctgaaaca tgagctgtgt 60
 ccactcaggg ttaaattgat taagggcggt gcaagatgtg ctttgtaaa cagatgcttg 120
 aaggcagcat gc 132

<210> 93
 <211> 275
 <212> DNA
 <213> HUMAN

<400> 93
 gcatagtgcc aggtgggtgcc ttgggttcca agctagtcca tggccccgat aaccttctgc 60
 ctgtgcacac acctgcccct cactccaccc ccatactagc tttggtatgg gggagagggc 120
 acagggccag acaaacctgt gagactttgg ctccatctct gcaaaagggc gctctgtgag 180
 tcagcctgct cccctccagg cttgctcctc ccccaccag ctctcgtttc caatgcacgt 240
 acagcccgtc cacaccgtgt gctgggacac cccac 275

<210> 94
 <211> 89
 <212> DNA
 <213> HUMAN

<400> 94
 ctgtccccct ccaggcttgc tcttccccc cccagctctc gtttccaatg cacgtacagc 60
 ccgtacacac cgtgtgctgg gacaccca 89

<210> 95
 <211> 61
 <212> DNA
 <213> HUMAN

<400> 95
 caccagctc tcgtttccaa tgcacgtaca gcccgtagc accgtgtgct gggacacccc 60
 a 61

<210> 96
<211> 116
<212> DNA
<213> HUMAN

<400> 96
acctgccct cactccaccc ccatacctagc ttggtatgg gggagagggc acagggccag 60
acaaacctgt gagactttgg ctccatctct gcaaaagggc gctctgtgag tcagcc 116

<210> 97
<211> 36
<212> PRT
<213> HUMAN

<400> 97
Gly Glu Glu Asp Leu Pro Ser Glu Glu Asp Ser Pro Arg Glu Glu Asp
1 5 10 15
Pro Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro Gly Glu
20 25 30
Glu Asp Leu Pro
35

<210> 98
<211> 6
<212> PRT
<213> HUMAN

<400> 98
Gly Glu Glu Asp Leu Pro
1 5

<210> 99
<211> 4
<212> PRT
<213> HUMAN

<400> 99
Glu Glu Asp Leu
1

<210> 100
<211> 5
<212> PRT
<213> HUMAN

<400> 100
Glu Glu Asp Leu Pro
1 5

<210> 101
<211> 6
<212> PRT
<213> HUMAN

<400> 101
Glu Asp Leu Pro Ser Glu
1 5

<210> 102
<211> 7
<212> PRT
<213> HUMAN

<400> 102
Glu Glu Asp Leu Pro Ser Glu
1 5

<210> 103
<211> 6
<212> PRT
<213> HUMAN

<400> 103
Asp Leu Pro Gly Glu Glu
1 5

<210> 104
<211> 22
<212> PRT
<213> HUMAN

<400> 104
Gly Gly Ser Ser Gly Glu Asp Asp Pro Leu Gly Glu Glu Asp Leu Pro
1 5 10 15

Ser Glu Glu Asp Ser Pro
20

<210> 105
<211> 25
<212> PRT
<213> HUMAN

<400> 105
Gly Glu Glu Asp Leu Pro Ser Glu Glu Asp Ser Pro Arg Glu Glu Asp
1 5 10 15

Pro Pro Gly Glu Glu Asp Leu Pro Gly
20 25

<210> 106

<211> 24
<212> PRT
<213> HUMAN

<400> 106
Glu Asp Pro Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro
1 5 10 15

Gly Glu Glu Asp Leu Pro Glu Val
20

<210> 107
<211> 7
<212> PRT
<213> HUMAN

<400> 107
Gly Glu Thr Arg Ala Pro Leu
1 5

<210> 108
<211> 7
<212> PRT
<213> HUMAN

<400> 108
Gly Glu Thr Arg Glu Pro Leu
1 5

<210> 109
<211> 7
<212> PRT
<213> HUMAN

<400> 109
Gly Gln Thr Arg Ser Pro Leu
1 5

<210> 110
<211> 1247
<212> DNA
<213> HUMAN

<220>
<221> misc_feature
<222> (1)..(1247)
<223> region 5' to the transcription initiation site as determined by RNase protection assay (nucleotide 3507 of SEQ ID NO: 5 and of Figures 2A-2F) in which an activating element is localized, which region corresponds to nucleotide 1328 to nucleotide 2574 of SEQ ID NO: 5 and of Figures 2A-2F.

<220>
<221> unsure what base is at position 647

<222> (647)

<223> unsure of the base at position 647, which is the same unknown base as that at position 1974 of SEQ ID NO: 5, and as that at position 1968 of SEQ ID NOS: 58 and 90. That unknown base at position 647 is in a region in which an activating element is localized and is 5' to the transcription initiation site.

<400> 110

```
tatgctactt tttgcagtcc tttcattaca tttttctctc ttcatttgaa gagcatgtta 60
tatcttttag cttcacttgg cttaaaaggt tctctcatta gcctaacaca gtgtcattgt 120
tggtagcact tggatcataa gtggaaaaac agtcaagaaa ttgcacagta atacttggtt 180
gtaagaggga tgattcaggt gaatctgaca ctaagaaact cccctacctg aggtctgaga 240
ttcctctgac attgctgtat ataggctttt cctttgacag cctgtgactg cggactattt 300
ttcttaagca agatagctta aagttttgtg agcctttttc cagagagagg tctcatatct 360
gcatcaagtg agaacatata atgtctgcat gtttccatat ttcaggaatg tttgcttgtg 420
ttttatgctt ttatatagac agggaaaactt gttcctcagt gacccaaaag aggtgggaat 480
tgttattgga tatcatcatt ggcccacgct ttctgacctt ggaaacaatt aagggttcat 540
aatctcaatt ctgtcagaat tggtagaaga aatagctgct atgtttcttg acattccact 600
tggtaggaaa taagaatgtg aaactcttca gttggtgtgt gtccctngtt tttttgcaat 660
ttccttctta ctgtgttaaa aaaaagtatg atcttgctct gagaggtgag gcattcttaa 720
tcatgatctt taaagatcaa taatataatc ctttcaagga ttatgtcttt attataataa 780
agataatttg tctttaacag aatcaataat ataatccctt aaaggattat atctttgctg 840
ggcgtagtgg ctacacactg taatcccagc actttgggtg gccaaagggtg aaggatcaaa 900
tttgcctact tctatattat cttctaaagc agaattcatc tctcttccct caatatgatg 960
atattgacag ggtttgccct cactcactag attgtgagct cctgctcagg gcaggtagcg 1020
ttttttgttt ttgtttttgt ttttcttttt tgagacaggg tcttgctctg tcaccaggc 1080
cagagtgcaa tggtagctgc tcagctcact gcagcctcaa ccgcctcggc tcaaaccatc 1140
atcccatttc agcctcctga gtagctggga ctacaggcac atgccattac acctggctaa 1200
tttttttgta tttctagtag agacagggtt tggccatgtt gcccggtg 1247
```

<210> 111

<211> 17

<212> DNA

<213> HUMAN

<400> 111

```
ctctgtgagt cagcctg 17
```

<210> 112

<211> 23

<212> DNA

<213> HUMAN

<400> 112

```
aggcttgctc ctccccacc cag 23
```

<210> 113

<211> 18

<212> DNA

<213> HUMAN

<400> 113

```
agactttggc tccatctc 18
```


<210> 114
<211> 20
<212> DNA
<213> HUMAN

<400> 114
cactccaccc ccatacctagc 20

<210> 115
<211> 26
<212> DNA
<213> HUMAN

<400> 115
gggagagggc acagggccag acaaac 26

<210> 116
<211> 20
<212> PRT
<213> HUMAN

<400> 116
Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
1 5 10 15

Gly Gly Gly Ser
20